Emergence et Sémantique des Mondes Possibles pour des Agents Appris

Joël Quinqueton*

- LIRMM Montpellier

jq@lirmm.fr

RESUME : Ce travail se place dans le cadre d'une approche émergentiste des systèmes multi-agents. Nous nous intéressons en particulier à l'émergence de concept dans une organisation d'agents appris. Le concept peut alors être défini comme une entité qui se renforce à l'usage, un peu comme les tas dans une termitière artificielle. La sémantique des mondes possibles permet selon nous de proposer un cadre formel bien adapté à cette vision multi-agent d'un processus d'apprentissage.

1 Introduction

Une approche traditionnelle en Intelligence Artificielle est de proposer des méthodes d'apprentissage en s'inspirant des stratégies qu'ont développé les systèmes naturels. Ces méthodes reposent le plus souvent sur une formalisation *a priori* du domaine du problème. Dans les domaines distribués et évolutifs, cette formalisation a priori devient difficile et requiert une capacité adaptative accrue. La nature semble avoir résolu ce problème par émergence de comportements collectifs qui s'auto-organisent à partir de la dynamique même des entités, par interactions entre eux et avec l'environnement.

Mieux comprendre et exploiter ces phénomènes, dits émergents, doit alors fournir non seulement une approche alternative pour la modélisation informatique, mais également pour la résolution de problèmes. Cette approche, dont le principe est de construire un collectif d'agents, plongé dans un environnement, qui de par leurs interactions vont évoluer ensemble jusqu'à converger vers un état stable représentatif d'une solution, a été initiée par R.Brooks en Robotique [Brooks 89, Brooks 91].

Nous nous proposons ici d'étudier les méthodes d'apprentissage automatique sous l'angle des communications entre des agents d'un système émergent.

Après avoir rappelé les principales définitions relatives à la sémantique des mondes possibles et quelques résultats de travaux récents sur l'émergence, nous proposerons une définition non complètement classique de l'apprentissage automatique et nous présenterons notre approche de l'étude des échanges de messages par l'environnement dans des processus d'apprentissage.

2 Une extension de la sémantique des mondes possibles

La sémantique des mondes possibles est un formalisme classique permettant de représenter, dans le cadre des logiques modales, les croyances d'un agent. Si l'on représente dans ce cadre les croyances d'un ensemble d'agents, il sera possible de représenter leurs échanges de connaissances.

Nous pensons que ce formalisme est particulièrement intéressant pour l'apprentissage automatique [Holland 89, Mitchell 97], car les exemples peuvent être considérés comme des mondes connus (c'est à dire sur lesquels on connait le concept). La généralisation peut alors être considérée comme l'ensemble des mondes depuis lesquels un exemple au moins est accessible.

Cela permet de définir l'apprentissage d'une manière adaptée au contexte multi-agent. Nous reviendrons plus loin sur ce point

2.1 Rappels sur les systèmes propositionnels modaux

Un système formel se définit habituellement par une syntaxe (ou langage), définissant la notion de formule bien formée, des axiomes et des règles de dérivation.

2.1.1 Les formules bien formées

On suppose donné un ensemble L_0 de symboles propositionnels. Par définition, les éléments de cet ensemble sont des formules bien formées (fbf).

Un élément singulier de L_0 est le symbole $\bot$ représentant une proposition toujours fausse (absurde).

Si f et g sont des fbf, alors $f \Rightarrow g$ est une fbf. Cette opération est appelée implication (parfois implication matérielle).

Par définition, on définit les symboles suivants :

- la négation $\neg f \equiv f \Rightarrow \bot$,
- la disjonction $f \vee g \equiv \neg f \Rightarrow g$,
- la conjonction $f \wedge g \equiv \neg(\neg f \vee \neg g)$.

Si on suppose donné un ensemble $A = \{a_i, i = 1, n\}$, dont les éléments seront appelés *agents*, alors $c(a_i, f)$, qui se lit *l'agent a_i croit que f* est une fbf pour toute fbf *f*.

2.1.2 Les axiomes

Par définition, les axiomes sont un ensemble initial donné de fbf qui pourront être interprétées comme des formules toujours vraies. Il existe plusieurs axiomatiques de

la logique propositionnelle, qui ont été proposées dans la littérature. Un ensemble d'axiomes est minimal si aucun de ses éléments ne peut être dérivé à partir des autres, en utilisant les règles de dérivation du système formel.

Plusieurs systèmes d'axiomes sont possibles pour définir la logique propositionnelle. Certains sont composés de 4 axiomes, définis à partir de la disjonction. Un autre schéma, proposé par L. Catach [Catach 90] est composé de 3 axiomes :

(A1) $g \Rightarrow (f \Rightarrow g)$

(A2) $(f \Rightarrow (g \Rightarrow h)) \Rightarrow ((f \Rightarrow g) \Rightarrow (f \Rightarrow h))$

(A3) $(\neg f \Rightarrow \neg g) \Rightarrow ((\neg f \Rightarrow g) \Rightarrow f)$

Un autre schéma de 3 axiomes est proposé par Laurière [Laurière 86], qui est identique au précédent sauf pour le 3ème axiome :

(A3) $(\neg f \Rightarrow \neg g) \Rightarrow (g \Rightarrow f)$

Il est possible de démontrer que ces schémas d'axiomes sont minimaux.

Si on suppose défini un ensemble d'agents, alors on ajoute habituellement l'axiome de distribution :

(K) $c(a_i, f \Rightarrow g) \Rightarrow (c(a_i, f) \Rightarrow c(a_i, g))$

2.1.3 Les règles de dérivation

Elles sont au nombre de 2 pour la logique propositionnelle :

la règle de substitution (R1)

si f est un théorème contenant le symbole p et si g est une fbf, alors la fbf obtenue par substitution de g à p dans f est un théorème.

la règle du *Modus Ponens* (R2)

si f et $f \Rightarrow g$ sont deux théorèmes alors g est un théorème.

On peut, à partir de ces deux règles, montrer par exemple que ces systèmes formels vérifient (c'est à dire admettent comme théorèmes) les 3 principes d'Aristote :

Identité : $f \Rightarrow f$

Tiers exclu : $f \vee \neg f$

Non contradiction : $\neg(f \wedge \neg f)$

Si on suppose défini un ensemble d'agents, alors on ajoute habituellement :

la règle de nécessitation (R3)

si f est un théorème alors $c(a_i, f)$ est un théorème.

Mais nous verrons que cette règle n'est pas toujours souhaitable dans le cas de raisonnements multi-agents, et qu'elle est caractéristique d'un type d'agent particulier.

2.2 La sémantique des mondes possibles

Cette sémantique a été introduite par Kripke [6], et beaucoup de travaux ont développé ce domaine. Elle est intéressante pour l'apprentissage, car les exemples peuvent être considérés, en un certain sens, comme des mondes possibles particuliers. Nous commencerons donc par présenter ce formalisme.

2.2.1 *Définitions*

Un modèle M est fait d'un ensemble de mondes W, un ensemble d'agents A et un ensemble de relations binaires entre ces mondes R_i, un langage formel L de formules tel que nous l'avons défini précédemment (qui produit un ensemble $Fma(\Box)$), et une fonction d'évaluation I qui donne, pour chaque formule de $\Box$ et chaque monde une valeur dans l'ensemble *{true, false}*.

Chaque relation binaire R_i est liée à un agent et est appelée sa *relation d'accessibilité* entre mondes.

Un tel modèle est noté $M = \langle W, L, R_1, \ldots, R_i, \ldots, I \rangle$.

La fonction I est telle que $I(f \Box g, w) = true \equiv I(g, w) = true$ or $I(f, w) = false$.

Cette propriété garantit la transitivité de l'opérateur d'implication, et la formule suivante est vraie dans tous les mondes: $(f \Box g) \Box ((g \Box h) \Box (f \Box h))$.

L'élément particulier $\Box$ de $Fma(\Box)$ vérifie $\Box w \Box W, I(\Box, w) = false$. L'opérateur de négation est donc tel que $I(\Box f, w) = I(f \Box \Box, w) = true \equiv I(f, w) = false$.

La croyance de chaque agent est définie à partir de sa relation d'accessibilité. Un agent a_i croit que f est vraie dans w si f est vraie dans tout monde auquel l'agent a_i a accès depuis w.

Formellement, $I(c(a_i, f), w) = true \equiv \Box x; wR_i x, I(f, x) = true$.

2.2.2 *Tautologies et axiomes*

Le principal intérêt de cette sémantique est que des axiomes additionnels peuvent être déduits des propriétés de la relation d'accessibilité.

Propriété de la Relation	**Tautologie**
réflexive $\Box w(wR_i w)$	connaissance (T) $c(a_i, f) \Box f$
faiblement réflexive $\Box w(xR_i w \Box wR_i w)$	confiance (U) $c(a_i, c(a_i, f) \Box f)$
symétrique $\Box w(wR_i x \Box xR_i w)$	(B) $f \Box c(a_i, \Box c(a_i, \Box f))$
serielle $\Box w \Box x(wR_i x)$	non contradiction (D) $c(a_i, f) \Box \Box c(a_i, \Box f)$
transitive $\Box w \Box x \Box y(wR_i x \Box xR_i y \Box wR_i y)$	introspection positive (4) $c(a_i, f) \Box c(a_i, c(a_i, f))$
Euclidienne $\Box w \Box x \Box y(wR_i x \Box wR_i y \Box xR_i y)$	introspection negative (5) $c(a_i, f) \Box c(a_i, \Box c(a_i, \Box f))$
pseudo-fonctionnelle $\Box w \Box x \Box y(wR_i x \Box wR_i y \Box x = y)$	(D_c) $\Box c(a_i, \Box f) \Box c(a_i, f)$
fonctionnelle stricte: serielle et pseudo-fonctionnelle	(D+D_c) $c(a_i, f) \Box \Box c(a_i, \Box f)$
dense $\Box w \Box x(wR_i x \Box \Box y(wR_i y \Box yR_i x))$	(4_c) $c(a_i, c(a_i, f)) \Box c(a_i, f)$
faiblement connexe $\Box w \Box x \Box y(wR_i x \Box wR_i y \Box xR_i y \quad yR_i x \quad x = y)$	(H) $(c(a_i, f \quad g) \Box c(a_i, c(a_i, f) \quad g) \Box c(a_i, f \quad c(a_i, g))) \Box c(a_i, f) \quad c(a_i, g)$
faiblement orientée $\Box w \Box x \Box y(wR_i x \Box wR_i y \Box \Box z(xR_i z \Box yR_i z))$	(G) $\Box c(a_i, \Box c(a_i, f)) \Box c(a_i, \Box c(a_i, \Box f))$

Le résultat fondamental dans ce domaine est le théorème de Goldblatt [Goldblatt 86]. Nous en rappelons ici les grandes lignes, dont nous avons besoin pour la suite.

Le théorème de Goldblatt énonce l'équivalence entre chacune des 11 propriétés de la relation d'accessibilité et chacune des 11 tautologies du système logique correspondant. Ces 11 cas sont résumés dans la table suivante.

Donc, dans tous ces cas, le théorème de Goldblatt pose que chaque schéma pour lequel la relation a une des propriétés possède la tautologie correspondante comme axiome. Et, réciproquement, tout schéma possédant une des tautologies a une relation d'accessibilité possédant la propriété correspondante.

3 Rappel de quelques définitions récentes de l'émergence

Nous rappelons ici les traits essentiels et les conclusions d'un travail sur l'émergence réalisé récemment dans le cadre d'un groupe de travail [MRJean 97].

Ce travail s'est appuyé, en ce qui concerne les définitions de l'émergence applicables aux SMA, sur 3 acceptions de l'émergence□

- la notion de calcul émergent (*Emergent computation*) de Stéphanie Forrest [Forrest 90],
- les 3 types d'émergence définis par Yves-Marie Visetti [Visetti 96] pour les sciences cognitives, ainsi que la notion de phénomène émergent de Marvin Minsky [Minsky 88],
- la notion d'émergence apparaissant en économie et en gestion.

Dans le domaine des systèmes multi-agents, il semble essentiel de se donner une définition positive, *temporelle* (où apparaît explicitement le temps) et *constructive* de l'émergence.

La première caractéristique essentielle d'un système multi-agent est qu'aucun agent ne contrôle complètement la dynamique de la population. Les agents sont *limités* et il y a des éléments du système global qu'ils ignorent. Il y a donc un extérieur relativement à chaque agent : un *environnement*. Cet aspect est important dans notre travail, puisqu'il fonde l'intérêt d'une résolution distribuée.

La seconde caractéristique est que, par définition, les agents agissent et donc modifient cet environnement. Mais, les agents ne peuvent percevoir ou agir que localement dans cet environnement. Autrement dit, chaque agent *interprète* l'environnement suivant ses moyens limités (d'après les distinctions qu'il peut faire).

La troisième caractéristique est que l'extérieur de chaque agent contient d'autres agents. Les agents sont plusieurs dans un environnement commun (ils sont extérieurs les uns par rapport aux autres). Les interprétations de l'environnement par les divers agents peuvent être différentes. Dans le cas des agents réactifs, l'environnement contient des objets et d'autres agents. Dans le cas des agents cognitifs, l'environnement peut aussi contenir des messages.

Ainsi, la dynamique du système procède d'une itération entre interprétation de leur environnement local par les agents, action des agents sur cet environnement, nouvelle interprétation de l'environnement modifié, nouvelles actions, etc.

Quand une telle dynamique (ou certaines de ses composantes) se stabilise on peut parler *d'émergence* d'une structure ou de fonctionnalité globale. En effet, dans ce cas c'est l'état global émergent stable qui conditionne (ou sélectionne) les comportements individuels de chaque agent.

Une définition plus opérationnelle, caractérisant le tout et les parties et surtout la rétroaction du tout sur les parties, s'inspire des définitions précédentes et suppose :

- un système d'entités en interactions dont l'expression des états et de la dynamique se fait dans un vocabulaire ou une théorie ;
- ce système produit un phénomène, que ce soit un processus, un état stable, ou un invariant, qui est forcément global puisque produit par le système ;
- l'observation de ce phénomène global, soit par un observateur, soit par le système lui même (on parlera dans ce dernier cas d'émergence forte).

Cette observation ne peut se faire qu'à deux conditions⊡

1) une *inscription* du phénomène d'une part
2) une *interprétation* de cette inscription par l'observateur ou par le système d'autre part

Si l'on cherche à utiliser ces définitions en SMA, on peut partir du cas d'une émergence observée par ailleurs.

En effet, la démarche classique de modélisation et d'explication en émergence va de l'observation des phénomènes naturels jusqu'à leur reproduction avec des systèmes artificiels. Dans certains cas, la modélisation se restreint à une attitude *interprétative* des phénomènes observés sans pouvoir de validation des hypothèses avancées. C'est le cas pour certains comportements collectifs observés dans les sociétés d'insectes (altruisme, ...) ou sociétés humaines, et correspond à la démarche *herméneutique*.

Jean Pierre Müller [Muller 98] a récemment proposé une approche émergente de la spécification d'un problème⊡

- La description formelle du phénomène global que le système multi-agents doit réaliser
- La projection de ce phénomène global sur la structure d'interactions au niveau microscopique pour déterminer l'identité des agents et la dynamique d'interaction
- la spécification des comportements individuels des agents pour produire les interactions engendrant le phénomène global que l'on veut pouvoir observer.

C'est une approche de ce type que nous nous efforçons d'illustrer sur les problèmes d'apprentissage, en nous appuyant sur la sémantique des mondes possibles.

4 La communication dans les systèmes émergents

Dans les systèmes multi-agents en général, la communication se fait selon différents modes, en fonction du type d'agent et des contraintes du problème.

Traditionnellement, on considère que les agents cognitifs communiquent directement entre eux (par envoi de message point à point), tandis que les agents réactifs communiquent toujours par l'intermédiaire de l'environnement. Cette différence est à rapprocher des modèles de machines parallèles à passation de message ou à mémoire partagée.

Par ailleurs, différents types de messages apparaissent en fonction de la manière de désigner leur destinataire. Il existe une manière "opportuniste" qui consiste à ce que le destinataire soit uniquement le premier agent qui "prend" le message : c'est le cas, par exemple, des fourmis qui se transmettent des messages sous forme de phéromone. On peut parler dans ce cas de message "situé" car un tel message n'a aucun intérêt pour un agent qui n'est pas à l'endroit où a été déposé le message.

Nous illustrons ces idées sur de petits exemples, réalisés à l'aide de la plate-forme Starlogo[1].

4.1 Exemple des termites

Le premier exemple est inspiré par le comportement des termites rassemblant des copeaux de bois en piles. Elles suivent des règles simples, et constituent un bon exemple de tâche "pure" (sans communication ni but explicite).

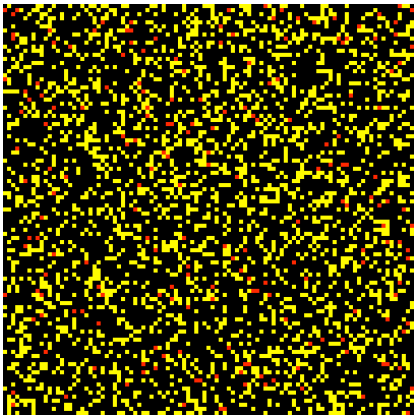
Chaque termite commence à chercher au hasard. Si elle rencontre un copeau de bois, elle le prend, et continue de se déplacer au hasard. Si elle rencontre encore un copeau de bois, elle recherche une place libre à coté de lui et pose son copeau de bois.

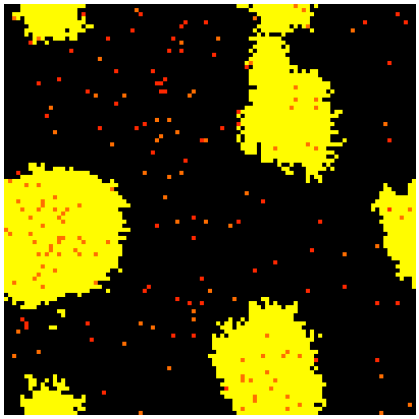
Avec ces simples règles, les copeaux finissent par se retrouver dans une pile unique. Au fur et à mesure que les piles commencent à se former, elles ne sont protégées en aucune façon. C'est à dire que les termites prennent quelquefois des copeaux dans des piles déjà constituées. Cette stratégie peut paraître contre productive, mais si les piles étaient "protégées", cela se terminerait par de nombreuses petites piles, pas par une seule grande.

En général, le nombre de piles décroît avec le temps. La raison en est simplement que des piles peuvent disparaître, lorsque des termites ont emporté tous les copeaux qu'elles contenaient, mais il n'y a aucun moyen qu'une nouvelle pile se forme à partir de rien, car les termites posent toujours leur copeau à coté d'un autre.

---

[1] Distribuée gratuitement par le MIT Media Lab





Le nombre de piles va donc aller en décroissant, car le seul moyen qu'il a d'augmenter est de couper une pile existante en 2. Par ailleurs, l'état dans lequel il ne reste qu'une seule pile est stable et constitue donc un invariant du système. Ceci est un bon exemple de stratégie décentralisée. Aucune termite n'est chargée d'une tâche particulière, et aucun endroit n'est désigné à l'avance pour les piles. Chaque termite suit un ensemble de règles simples, mais la colonie dans son ensemble accomplit une tâche relativement sophistiquée[2].

4.2 La causalité descendante et l'Emergence

Dans les définitions précédentes de l'émergence, nous avons montré le rôle de l'observateur pour caractériser un processus comme émergent. Un autre aspect important est la structure à plusieurs niveaux d'un système qui produit de l'émergence. Le niveau supérieur est posé par l'observateur, comme la notion de pile dans l'exemple des termites.

L'influence causale entre 2 tels niveaux d'abstraction dans l'émergence est du niveau supérieur vers le niveau inférieur, c'est à dire de façon descendante, et non ascendante, comme on pourrait le croire intuitivement. Dans l'exemple des termites, on constate cette causalité dans la manière d'utiliser le concept de pile pour expliquer pourquoi le nombre de piles est décroissant en probabilité.

Notre but est d'essayer d'utiliser ce schéma comme mécanisme de décision dans un ensemble d'agents appris. Commençons donc par définir ce qu'est un agent appris.

5 Agents appris

Une définition générale et consensuelle de l'apprentissage [Mitchell 97] consiste à poser qu'il y a apprentissage ou qu'un programme apprend :

- par l'expérience E par rapport à une classe de tâches T et une mesure de performances P
- si sa performance pour les tâches de T mesurée par P s'améliore avec E.

---

[2]Notre "cyber-termitière" peut être vue comme une machine de Turing avec une "tête de lecture" à déplacement aléatoire sur un "ruban" bidimensionnel.

L'expérience est constituée par un ensemble d'exemples, et la tâche est ici de généraliser ces exemples, c'est à dire de reconnaître un concept C, commun aux exemples mais inconnu au départ et extrapolable à d'autres objets.

Nous commençons par introduire une formule particulière qui signifie "être un exemple du concept C" et qui est vraie sur chaque monde représentant un exemple, et fausse dans les autres. Nous posons donc comme axiome que le concept C est vrai au moins dans chaque monde qui est un exemple de C.

Nous décrivons finalement, en partant de leur relation d'accessibilité, des cas particuliers d'agents jouant un rôle dans le processus d'apprentissage.

5.1 L'oracle, la sonde, le maître

Si la relation R_i est vide, alors tous les mondes sont isolés les uns des autres, et toutes les tautologies sont vraies, sauf (T) et (D), car une relation vide a ttoutes les propriétés sauf la réflexivité et la sérialité. L'agent correspondant croit tout et son contraire. Un tel agent est appelé **sonde**.

Si la relation R_i est réduite aux boucles autour des exemples $wR_ix \equiv (w = x)$, alors elle a toutes les propriétés, et toutes les tautologies sont vraies.

Un tel agent vérifie la règle de nécessitation, c'est à dire qu'il croit à tous les théorèmes: si f est un théorème alors $c(a_i, f)$ est un théorème. De plus, cet agent vérifie l'axiome de connaissance, c'est à dire que tout ce qu'il croit est vrai, et n'a pas de contradiction. Un tel agent est appelé un **oracle**.

Si la relation R_i est universelle, alors elle vérifie toutes les propriétés, sauf la pseudo-fonctionnalité. Donc, toutes les tautologies sauf (D_c) sont vérifiées. Un tel agent est appelé un **maître**.

Ces 3 types d'agents ont un rôle dans le processus d'apprentissage car ils représentent 3 types d'interactions d'un agent apprenant avec son environnement.

5.2 Un exemple d'apprenti

Soit $e \in E \subseteq W$ un exemple, c'est à dire un monde particulier interprétant le langage formel L. E sera par définition l'ensemble d'exemples. L'interprétation I est donnée, et g est le concept à apprendre.

Supposons que notre agent ait appris une formule $f \in L$ comme bonne description de g. La relation d'accessibilité est alors définie par:

$$f^+ = \{w \in W; I(f, w) = true\}$$

$$xR_f y \equiv (x, y) \in f^+ \cup (E - f^+)$$

Comme g est vraie sur les exemples, on obtient: $\forall w \in f^+, I(c(a_f, g), w) = true$, ce qui signifie que, si f est vraie, alors notre agent croit que le concept g est vrai.

Les axiomes vérifiés par notre agent sont alors:

(U) $c(a_f, c(a_f, g) \rightarrow g)$

(4) $c(a_f, g) \rightarrow c(a_f, c(a_f, g))$

(5) $c(a_f, g) \rightarrow c(a_f, \neg c(a_f, \neg g))$

D'après la définition habituelle de l'implication donnée précédemment, notre agent est tel que $c(a_f, f \rightarrow g)$ est toujours vraie, ce qui est bien le but de l'étape d'apprentissage.

Pour obtenir la non contradiction, la relation doit être sérielle, et pour cela chaque monde doit pouvoir accéder à au moins un autre. Pour avoir une telle situation, nous avons besoin d'exemples négatifs.

Soient donc 2 ensembles d'exemples, positifs et négatifs: $E^+, E^- \subseteq W$.

On définit comme précédemment :

$f^+ = \{w \in W; I(f, w) = true\}$

$f^- = \{w \in W; I(f, w) = false\}$

$xR_f y \equiv (x, y) \in f^+ \times (E^+ \cap f^+) \cup f^- \times (E^- \cap f^-)$

La relation est donc sérielle si $f^+ \cup f^- = W$, $E^+ \cap f^+ \neq \varnothing$ et $E^- \cap f^- \neq \varnothing$ c'est à dire si la formule *f* est évaluable dans tout monde et si elle est vraie sur au moins un exemple positif et fausse sur au moins un exemple négatif.

5.3 Prise de décision par coordination entre agents appris

Nous proposons un schéma d'apprentissage permettant d'utiliser les définitions précédentes pour généraliser un ensemble d'exemples décrits par un ensemble de variables booléennes.

Nous considérons successivement chaque variable comme le concept à apprendre, qui est connu cette fois. Chaque variable est utilisée pour diviser l'ensemble d'exemples en 2, les positifs et les négatifs, et l'on a : $E^+ = E \cap g^+, E^- = E \cap g^-$.

Nous pouvons alors appliquer la méthode précédente pour apprendre cette variable à partir des autres, et nous obtenons un ensemble de règles permettant de prédire cette variable à partir des autres.

La relation d'accessibilité est définie par :

$xR_f y \equiv (x, y) \in f^+ \times (E \cap g^+ \cap f^+) \cup f^- \times (E \cap g^- \cap f^-)$

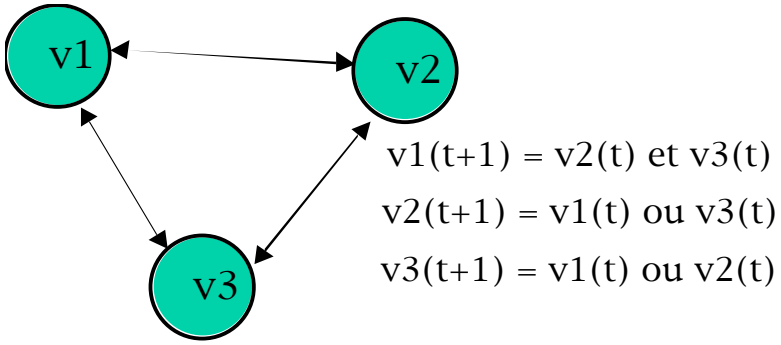
Nous obtenons donc pour notre ensemble de variables, un ensemble de règles, dont chacune prédit une variable à partir des autres. Nous obtenons ainsi un réseau booléen.

Un tel réseau [Kauffman 93] est constitué de N variables booléennes liées par un réseau de dépendances et des règles de transition. Ces réseaux produisent des états cycliques (Etats stables, Cycles et Attracteurs) au sein desquels restent les vecteurs booléens d'un pas de temps au suivant. Chaque règle agit comme une "termite" qui modifie son environnement en "corrigeant" la description courante.

L'étude de grands réseaux de ce type met en évidence l'importance du paramètre K représentant le nombre d'entrées de chaque élément (demi degré entrant). L'étude de la sensibilité aux perturbations (stabilité homéostatique) en fonction de ce paramètre K donne les résultats suivants [Kauffman 93] :

- Lorsque K décroît de N à 1, transition de phase pour K = 2
- Comportement chaotique pour K > 2

Du point de vue de l'homogénéité, le paramètre important est P, qui vaut toujours entre 1/2 et 1, et qui mesure la proportion maximum de valeurs identiques (1 ou 0) dans les sorties des règles de transition. Par exemple, P=1 correspond à une tautologie ou à une inconsistance (toujours faux).



Les propriétés des réseaux booléens sont alors les suivantes□

- Longueur des cycles: en √N pour K= 1 ou 2, exponentielle pour K grand
- Nombre d'attracteurs□exponentiel si K=1, √N si K=2, linéaire pour K grand (croît comme -logP)
- Stabilité homéostatique : faible sauf pour K=2
- Atteignabilité□élevée sauf pour K=2

Avec les définitions précédentes, un agent appris peut être considéré comme un agent capable de corriger un nouvel exemple. Cette correction est effectuée par le réseau booléen des agents. Il peut être considéré comme une règle de cohérence *faible* et *adaptative*.

Donc, comme dans l'exemple des termites, le concept est dans ce cas assez stable, car il a plus de probabilité de rester tel quel que d'être modifié.

5.4 Discussion

En utilisant ce formalisme, nous avons donc pu spécifier un type d'agent apprenant qui donne naturellement un mode d'emploi de la connaissance apprise, en particulier au travers des propriétés (à partir des axiomes) de l'agent appris. Bien sûr, d'autres cas peuvent être représentés en utilisant ce schéma.

En introduisant les différents types d'agents impliqués dans un processus d'apprentissage, nous pouvons définir avec précision la sémantique de chaque partie du processus d'apprentissage.

Nous pensons que l'intérêt d'une telle formalisation de l'apprentissage est de construire un système dans lequel l'utilisation de la connaissance apprise peut être formulée comme l'établissement d'une communication entre des agents à travers l'environnement. Dans ce processus de communication, la manière pour un agent de gérer les croyances d'un autre peut être définie axiomatiquement de façon à obtenir le résultat désiré. Ce modèle peut être rapproché de travaux d'Intelligence Artificielle récents comme le raisonnement à ressources limitées [Koriche 97], la logique des

choix indépendants [Poole 97, Poole 98], et de certains travaux autour des actes du langage [Perrault 90].

Le modèle des automates cellulaires est un cas particulier classique de réseau booléen, très utilisé dans l'étude des systèmes distribués. Notre formalisme est donc un modèle d'apprentissage d'automates cellulaires.

6 Conclusion et perspectives

Nous avons essayé ici de caractériser un phénomène émergent par la communication entre les agents du système qui le mettent en œuvre. Il apparaît clairement que le phénomène émergent proprement dit se produit après que la communication ait permis d'établir les rôles de chacun.

Il reste quelques points importants à approfondir pour rendre pleinement opératoire une approche de ce type⊡

- Nous étudions également l'utilisation des protocoles de coordination autoritaire [Segal 98, Segal 99] pour produire la décision finale.
- Il faut également introduire une fonction de coût pour évaluer les modifications de la description traitée.
- La détection de la stabilité (terminaison) par l'utilisateur (émergence faible) ou par les agents (émergence forte) reste encore un problème ouvert.

Nous pensons que ce travail constitue le début d'un cadre général d'étude des processus émergents.

L'étude des relations spatiales et temporelles dans l'occupation des territoires en Géographie est une application du travail décrit ici. Pour ce type d'application, la sémantique des mondes possibles est particulièrement bien adaptée, car les mondes possibles sont ici les hypothèses pouvant être faites sur l'usage future d'une parcelle de territoire.

7 Références

[BROOKS 89] R.A. BROOKS, A Robot that Walks: Emergent Behaviors From a Carefully Evolved Network, Massachusetts Institute of Technology, Cambridge, MA, A.I Memo. 1091, 1989.

[BROOKS 91] R.A.BROOKS, Intelligence Without Reason, IJCAI'91, 1991.

[CATACH 90] L. CATACH, Logiques non classiques. Présentation et applications à l'Intelligence Artificielle , Hermès (1990)pp. , 19-58, Chapter 1.

[CHELLAS 80] B.F. CHELLAS, *Modal Logic: an introduction*, Cambridge University Press (1980).

[FERBER 95] J. FERBER, *Les systèmes multi-agents*, InterEditions, 1995.

[FORREST 90] S. FORREST, Emergent computation: Self-organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks, introduction to the proceedings of the ninth annual CNLS Conference, in *Emergent computation*, MIT Press, Cambridge, MA, 1991, pp 1-11.

[GIROUX 96] S. GIROUX, P. MARCENAC, J. QUINQUETON, AND J. R. GRASSO. Modelling and simulating self-organised critical systems. In European Simulation Multiconference proceedings, Budapest, H, june 1996. SCS, SiE.

[GOLDBLATT 86] R. GOLDBLATT, Logics of Time and Computation, CSLI - Center for the Study of Language and Information, Stanford University , CSLI Lecture Notes(1986).

[HAMADI 96] Y. HAMADI, Distribution de Gsat, Journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1996, Hermès, pp 189-199.

[HAMADI 99] J. QUINQUETON, Y. HAMADI, Communication et Emergence□ une épidémie chez les termites, Journées françaises sur les systèmes multi-agents, JFIADSMA, Novembre 1999, Hermès, pp 225-235.

[HOLLAND 89] J.H. HOLLAND, K.J. Holyoak, R.E. Nisbett, P.R. Thagard, *Induction□ Processes of Inference, Learning and discovery*, The MIT Press, 1989.

[KAUFFMAN 93] S.A. KAUFFMAN, *The Origins of Order□ Self Organization and Selection in Evolution*, Oxford University Press, 1993.

[KORICHE 97] F. KORICHE, Raisonnement approximatif dans les systèmes à base de connaissances coopératifs, Ph.D. dissertation, Montpellier II University — Sciences et Techniques du Languedoc, February 1997.

[KRIPKE 63] S.A. KRIPKE, Semantic analysis of modal logic I: normal propositional calculi. Zeit. Math. Logik Grund. Math. 9(1963), 67-96.

[LAURIERE 86] J.L. LAURIERE, Intelligence Artificielle: résolution de problèmes par l'homme et la machine, Eyrolles (1986).

[MINSKY 88] M. MINSKY, *La société de l'esprit*, Intereditions, 1988.

[MITCHELL 97] T. MITCHELL, *Machine Learning*, Mac Graw Hill, 1997.

[MRJEAN 97] M.R. JEAN, Emergence et SMA, Journées françaises sur les systèmes multi-agents, JFIADSMA, Avril 1997, Hermès, pp 323-342

[MÜLLER 98] J.P. MÜLLER, "□Vers une méthodologie de conception de systèmes de résolution de problèmes par émergence ", Journées françaises sur les systèmes multi-agents, JFIADSMA, Octobre 1998, Hermès, pp 355-371

[PERRAULT 90] C.R. PERRAULT, An Application of Default Logic to Speech Act Theory, MIT Press (1990), Chapter 9.

[POOLE 97] D. POOLE, The independent choice logic for modelling multiple agents under uncertainty. Artificial Intelligence (1997), accepted paper 1461.

[POOLE 98] D. POOLE, A.MACKWORTH, R. GOEBEL, *Computational Intelligence□ a Logical approach*, Oxford University Press, 1998.

[SEGAL 98] I. SEGAL, Communication complexity and Coordination by Authority, Working paper, June 3, 1998 draft, available from http://elsa.Berkeley.EDU/users/segal/

[SEGAL 99] I. SEGAL, Coordination and Discrimination in Contracting with externalities: Divide and Conquer?, Working paper, July 6, 1999 draft, available from http://elsa.Berkeley.EDU/users/segal/

[VISETTI 96] Y.M. VISETTI, Emergence et hétérogénéité : le problème des passages, in Journées thématiques de l'ARC (émergence et explication), 1996.